بعض الطرائق الحصينة لتقدير قدرة الطيف وفق نموذج ARMA "دراسة تجريبية"

أ.د. عبد المجيد حمزة الناصر / كلية الادارة والاقتصاد/ جامعة بغداد
الباحث/ سحر طارق محمود

تاريخ التقديم:2016/8/9
تاريخ القبول:2016/10/17

المستخلص :

يمكن تعريف "الاحصاءات الحصينة "المقاومة للاخطاء الناتجة من الانحراف عن فروض نظرية استقرارية العمليات الاحصائية (Asymptotically ، Approximately Met ،Reasonable UNBIASED، Reasonably Small Bias ، Efficient) في بيانات اختيرت من مدى واسع من التوزيعات الاحتمالية تتبع التوزيع الطبيعي او خليط من توزيعات اخرى بانحرافات معيارية مختلفة. وتؤدي دالة قدرة الطيف (power spectrum) دورا رئيسا في تحليل عمليات عشوائية مستقرة مرتبة وفق الزمن قد تكون منفصلة او مستمرة، ووصف قياس القدرة الاجمالية لها دالة في التردد.

تشترك طرائق التقدير الحصينة مع مفهوم اللامعلمية في عدم وجود نموذج واضح المعالم (Free distribution) اذ يتوزع جزء منه وفق التوزيع الطبيعي ، والاخر توزيعه غير معلوم فيصبح توزيعا ملوثا معلماته غير معلومة ، لذا يمكن اعتبار الطرائق الحصينة اعلى مستوى في درجات الطرق اللامعلمية تستند على تحويل احصاءة الاختبار الى درجة معيارية تجرى عليها عملية التقارب بالدرجة المعيارية، وبمقارنة بعض طرائق التقدير الحصينة واللامعلمية ، من خلال ايجاد افضل مقدر لقدرة الطيف وفق الانموذج المختلط ARMA لسلسلة تتبع توزيعا معينا، وتطبيقها على بيانات واقعية تعكس مصادر التغير في سلسلة متمثلة بالإشارة الكهربائية الـ (ECG - Electrocardiogram) المسجلة بحالة الراحة (غير المرضية) ، وذلك بواسطة تحليل قوة الطيف (Power Spectrum)، فكان الأسلوب المتبع في تنقية أشارة الضوضاء البيضاء بتقييد معلمتي الموضع والقياس باعتماد التقديرات الأولية الناتجة عن كافة التجارب بتطبيق طرائق التقدير الحصينة للمعلمتين على الإشارة الحقيقية في الحصول على نتائج واقعية ترتبط بالنظام الحقيقي لدراسة سلوك الإشارة الكهربائية (ECG) وذلك نتيجة لاحتواء قيم الإشارة على الملوثات بالقيم المضافة، وتبين ان مقدرات القدرة عند التوزيع الطبيعي تزداد بازدياد حجم المقطع ، وان واعلى مقدر طيف القدرة بطريقة Tukey عند المقطع 200 ، ثم طريقة Huber، Hampel، Andrew على التوالي.

المصطلحات الرئيسة للبحث/ الحصانة، قدرة الطيف، الانموذج المختلط،(الانحدار الذاتي،المتوسطات المتحركة) ، المحاكاة .

مجلة العلوم الاقتصادية والإدارية
العدد 97 المجلد 23
الصفحات 358-377

*البحث مستل من اطروحة دكتوراه

المبحث الأول

1-1 : **المقدمة**:

تتبع الكثير من بيانات العلوم التطبيقية في سلوكها نمط السلاسل الزمنية $X_1, X_2, \dots, X_n$ بتطورها عبر الزمن يوميا ، اسبوعيا، او شهريا،...والتي تمثل احدى العمليات العشوائية المستمرة $X_t^{(w)}$ ، والتحليل الاحصائي الدقيق المعني بتحليل سلوك الظواهر يقتضي توفر بيانات دقيقة للظاهرة بصورة نماذج تعتمد مشاهداتها على بعضها تمكن اجراء التحليل والتنبؤ المستقبلي بما تؤول اليه الظاهرة وفق أساليب رياضية رصينة يسمى بتحليل السلاسل الزمنية.

وحيث ان ليس سهلا توفير بيانات خالية من الشواذ (الشوارد) والتي هي " قيم خارجة عن النمط العام للبيانات ما يجعلها غير ملائمة لبعض او لجميع فرضيات الساسية النماذج الاحصائية (كالطبيعية، وصفتي الانعكاسية والإستقرارية) " والتي قد تؤثر بشكل مباشر وفعّال في تشكيل نموذج عملية التحليل، او قد تنحرف مشاهداتها عن النسق العام للبيانات مما يتعذر الحصول على مقدرات تتصف بالدقة والحصانة اذ ان وجود قيمة واحدة بعيدة عن نسق البيانات قد تفقد المقدرات أهميتها وخصائصها الجيدة.

تنشأ شواذ عادة عن توزيعات ثقيلة الاطراف **Heavy tailed distributions** او من اخطاء المشاهدة ، او من توزيعات مختلطة **mixture distributions** قد تؤثر في المقدرات وفق الطرائق الاعتيادية (طريقة الامكان الاعظم **Maximum Likelihood** التي تعتمد على معرفة التوزيع سوآءا كان للخطأ او للسعة والتي هي من أكثر طرائق التقدير دقة ومناسبة عند توافر شروط نماذج السلاسل الزمنية ، وطريقة المربعات الصغرى **Least Squares** ، وطريقة العزوم **Moment method** ... الخ).

2-1 : **الهدف من البحث** : مقارنة بعض طرائق التقدير الحصينة واللامعلمية ، من خلال ايجاد افضل مقدر لقدرة الطيف وفق الانموذج المختلط ARMA لسلسلة تتبع توزيعا معينا. وتطبيقها على بيانات واقعية تعكس مصادر التغير في سلسلة متمثلة بالإشارة الكهربائية الـ (ECG - Electrocardiogram) المسجلة بحالة الراحة (غير المرضية) في حالة الراحة، من خلال تحليل قوة الطيف (Power Spectrum).

3-1 : **الاستعراض المرجعي**:

- في عام 1927s عرض الباحث **Yule** نموذج الانحدار الذاتي بالرتب الاولى والثانية **AR(1,0)** ، **AR(2,0)** ووسعه **Walker** في عام 1931 سماه بالانحدار الذاتي العام **AR(p)**، وعرض **Stutsky** نموذج المتوسطات المتحركة **MA** في عام 1937، وفي عام 1954 عرض **Would** النموذج المختلط **ARMA** وبإدخال(AR) و (MA) تحول الاهتمام بعيدا عن تحليل مجال تردد السلاسل الزمنية الى مجال الزمن .[2].
- في عام 1930 بلغت الذروة في التحليل الطيفي العام بأطروحة **Winner Norbert Generalized Harmonic Analysis**"[5]، كان هذا التطور نقطة تحول في مجال تحليلات القدرة الرئيسة ، ورائدا في النظريات الحديثة للعمليات الاحتمالية اذ غطت هذه النظرية كلتا السلاسل الزمنية وحيدة المتغيرات والمتعددة، وطبقت على سلاسل احتمالية وغير احتمالية.
- في عام 1954 اكتشف **Would** [12] النماذج المختلطة **ARMA** ثم ابتدأت بعده العديد من الدراسات منها ساهمت باستنباط طرائق تقدير النماذج **AR** و**MA** و**ARMA** بكفاءة. والاخرى اهتمت بتعميم النتائج لتصل الى النماذج الموسمية.
- في عام 1962 عمم **Walker** [10] النتائج التي توصل اليها **Durbin** في عام 1960على نماذج **ARMA** وفي الفترة 1968- 1976 وضع كل من **(Box & Jenkins & Walks)** [8] أساليب مختلفة في تحليل السلاسل الزمنية والتنبؤ بها فيما يخص الانموذج المختلط **(ARMA)**.

- في عام 1997 قدم الباحث النقيب [2]، عبد الخالق عبد الجبار اطروحته الموسومة" التخمينات الحصينة لأطياف القدرة مع تطبيق عملي" تناول فيها عدة طرائق حصينة مقترحة لتقدير معلمات انموذج الانحدار الذاتي في سلسلة زمنية ممثلة بالإشارة الكهربائية للعضلة (EMG) وتعيين اي من المقدرات مناسبة لتقدير اطياف القدرة.
- في عام 2009 قدم كل من Nora Muller, Daniel Pena and Victor j. ,Yohai [9] تقدير-M الحصين للنماذج المختلطة ARMA وتوصلوا الى ان تأثير قيمة متطرفة واحدة يقتصر على الفترات الزمنية التي يحدث فيها، وانه يرتبط بشدة بتقديرات المرشح وقد استخدموا طريقة Monte Carlo لمقارنة التقديرات نسبة الى (تقديرات – M) القياسية واستند على اجراء التشخيص.

المبحث الثانـــي : الجانب النظري

1-2 : مقاييس الحصانة التقليدية:

- نقاط الانهيار Break Down Points

تمثل نسبة من البيانات الغير سليمة (ربما تكون كبيرة او صغيرة ويجب ان لا تتجاوز 50% من المشاهدات واذا حدث ذلك تعد البيانات ملوثة.[10]

- دالة التأثير : Influence Function
- منحنى الحساسية : Sensitivity

2-2 : معلمات التقدير الحصين : Parameters of Robust Estimation

- معلمات الموقع : Estimation of Location

المتوسط والوسيط حالات خاصة من (تقدير – M)، تقدر معلمة الموقع من خلال رسم توزيع Bootstrap .

- معلمات القياس : Estimation of Scale

يعد الانحراف المعياري تقدير القياس Scale التقليدي، يتأثر بشدة بوجود الشواذ بسبب تفاقم مربع الانحرافات عن المتوسط .[12]

3-2: التقديرات الحصينة لمعلمة الموقع: (Robust M-Estimates)

تُعتَمد طرائق التقدير الحصينة عموماً أسلوب التنعيم للقيم الحقيقية المتضمنة القيم الشاذة ومن النوع المعروف بالقيم المضافة (Additive Outliers) خاصة، حيث تنجم من خلال أحد ثلاث حالات هي أما أخطاء في القياس أو أخطاء المعاينة، أو لأسباب تتعلق بطبيعة الظاهرة المبحوثة.

تتعلق الحصانة بإمكانية التقدير الدقيق لمعلمات نموذج معين X_t في عملية ملوثة X_t^{ϵ} ، ولتحقيق حصانة التقدير يتم احتساب الخطأ (Residuals) باستخدام مرشحات حصينة مقترحة من قبل Masreliez والتي تنتهج اسلوب التنبؤ لخطوة واحدة في نماذج ARMA(P, q) بقيم شاذة مضافة AO ، عرفها Huber في عام 1964 بانها تقديرات تعد تعميم لتقديرات الامكان الاعظم ، وقريبة من تقديرات المربعات الصغرى عند عدم وجود شواذ فيها.[14]

تعرف (مقدرات-M) بعملية تصغير يكافئ تعظيم الامكان الاعظم MLE اذ يقوم بتعظيم $\prod_{i=1}^{n} f(x_i)$ وتصغير $\sum_{i=1}^{n} -log\, f\,(x_i)$ تتلخص باختيار الدالة العامة للنموذج الاحتمالي بحيث تعطي أدنى عدد من الشواذ ، ومن ثم بناء المقدر الحصين تلقائيا.[13]

وعموما فقد تم اختيار أربع طرائق حصينة للحصول على قيمة مقدر الموقع-M ، وهي كما يأتي:

2-3-1: **طريقة** (Huber) **الحصينة** :

اقترح Huber في عام 1964 دالة (ψ - Likelihood) محددة بتأثير المشاهدات y وفق الصيغة : $\frac{|y-\widehat{\mu}|}{\widehat{\sigma}}$ **تعرف بتقديرات الإمكان الأعظم الحصينة لتقدير معلمة النزعة المركزية – M، اذ ان اختلاف قيم الأوزان (ω_i) وفقا لطبيعة القيم المتطرفة، والتي تفترض عتبة قطع (c) تفصل ما بين مجموعتي القيم التي تحقق توزيعاً معلوماً والقيم المُلوَثة بالقيم الشاذة والتي تكون مجهولة التوزيع، وبموجب ذلك فأن القيم بالدرجات المعيارية ذات التوزيع المعلوم (الأقل من عتبة القطع) تكون قيمة الوزن المرجحة بها هو العدد (الواحد الصحيح)، في حين تكون قيم الأوزان المستخدمة متناقصاً بارتفاع القيم المعيارية، وبذلك تتضح آلية عمل الطريقة في تغيير القيم المعيارية الكبيرة وتقليل قيمة التغير بالدرجات المعيارية التي تقترب من المركز (الصفر).[10]**

يعرف مقدر الامكان الاعظم بالكفاءة التقاربية يمكن احتسابه عندما تكون الكثافة منتظمة، وقد اقترحت النماذج التالية من مقدرات الامكان (مقدرات – M)

$\hat{\alpha} = (\hat{\phi}', \hat{\theta}', \hat{\mu}')'$ **-: وفق الصيغة التالية** **[13]**

$$g(\alpha', \hat{\sigma}) = \sum_{i=P+1}^{n} Q[\frac{r_i(\alpha')}{\hat{\sigma}}] = \text{Min}$$

اذ ان : $r_i(\alpha')$: تشير الى البواقي

دالة Q : تشير الى تقييد تأثير الاخطاء الكبيرة في التقدير .

(α, σ) : معلمات (الموضع , والقياس) يمكن ايجادها معا من خلال حل المعادلة :

$$b = \frac{1}{n-2P-q-1} \sum_{i=P+1}^{n} \psi^2(\frac{r_i(\alpha')}{\sigma})$$

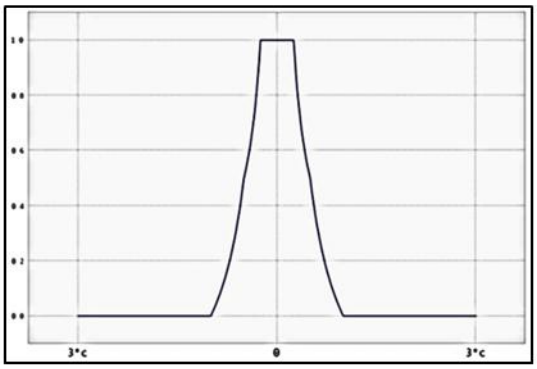
$$C = \frac{b}{2}(n-2P-q-1)$$

$$\psi(t) = \frac{d\,Q(t)}{dt}$$

$$\sum_{i=P+1}^{n} \psi(\frac{r_i(\hat{\alpha})}{\sigma^2})\, d_i^+(\hat{\alpha}) = 0$$

اذ ان : d_i^+ : يشير الى المشتقة الاولى للخطأ $r_i(\hat{\alpha})$.

تعرف الدالة (ψ -Likelihood) بـ (انحرافات - Ψ) في بيانات نموذج احصائي الغرض من هذه الدالة تقييد تأثير الاخطاء الكبيرة في التقدير لذا يجب ان تكون مقيدة ومستمرة يعبر عن تقدير ($\psi - Liklihood$) وفق الشكل التالي :

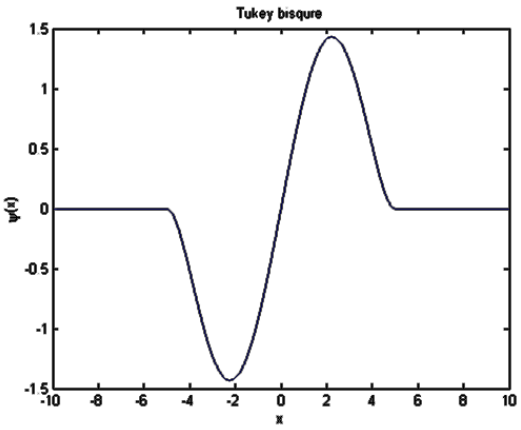


الشكل(1) : دالة تقدير الامكان(Function-ψ) لـ Huber [17]

2-3-2: **طريقة** (Tukey) **الحصينة**:

وتعرف بطريقة الوزن الثنائي بتخصيص الصفر للمشاهدات المعيارية التي تزيد عن المقدار (4.685)، وبتخصيص الأوزان بالتناسب العكسي مع المسافة من المركز ولكافة المشاهدات الأخرى.

(مقدرات-M) لـ Tukey تأخذ (Function-ψ) لأي قيمة C موجبة تعرف الصيغة وفق الجدول في ادناه :[18]



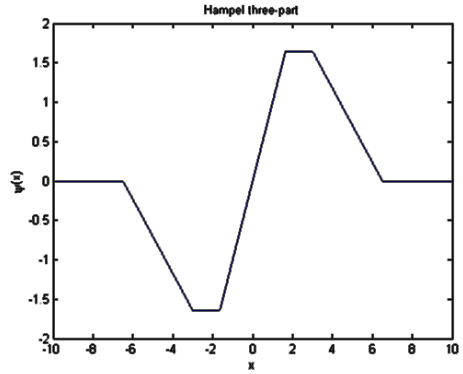
الشكل (2): دالة تقدير الامكان$(Function-\psi)$ لـ 'Tukey

3-3-2: **طريقة** (Hampel) **الحصينة**

$$W(x)=\frac{\psi(x)}{x}$$

$$\text{Min}\sum_i W(r_i^{(k-1)})r_i^2$$

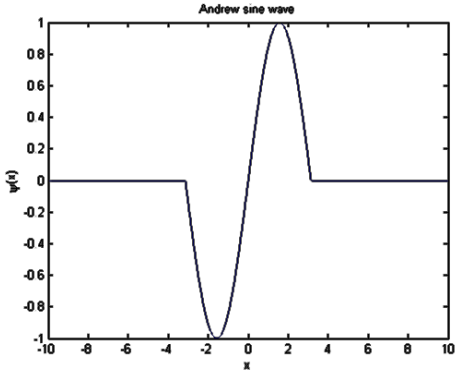
تقديرات Hampel الثلاث (تقديرات-M) تأخذ الدالة (Function-ψ) والتي هي دالة مفردة وتعرف بالصيغة الموضحة بالجدول في ادناه:[6]



الشكل (3) : دالة تقدير الامكان(Function-ψ) لـ 'Hampel

4-3-2: **طريقة** (Andrew) **الحصينة**:

هي طريقة مستخدمة بتطبيقات منحنى الجيب للأوزان بطريقة التنقية بالإلغاء وليس التعديل، وذلك من خلال عتبة القطع c التي يجرى بموجبها ضرب القيم المعيارية المطلقة بقيمة الصفر للدرجات المعيارية المرتبة والتي تزيد عن قيمة القطع المذكورة، والإبقاء على تلك الدرجات التي تقل عن العتبة المذكورة. وفي ضوء ذلك فأن القيم المعيارية لقيم الضوضاء البيضاء للإشارة ECG وللمقاطع الزمنية المختارة وهي:[14]
(TSV= 200 m s ،و TSV= 100 m s ،TSV= 50 m s ،TSV=25ms)
((لم تتجاوز جميع قيم الضوضاء البيضاء فيها قيمة عتبة القطع c=3))، الأمر الذي يتعين بموجبة اعتبار هذه الطريقة هي طريقة لامعلمية إضافة لكونها مصنفة من بين طرائق التقدير الحصينة لمعلمة الموقع – M، وذلك من خلال ارتفاع كفاءة التقدير تقاربياً (Asymptotically) بالدرجات المعيارية بزيادة حجم المقطع (العينة)، وهو ما تحقق فعلاً في الجانب التجريبي. الموجة الجيبية لـ (مقدرات – M) Andrew تأخذ (Function-ψ)



الشكل (4):دالة تقدير الامكان(Function-ψ) لـ Andrew [14]
من اشهر (مقدرات- M):

Type	$\rho(x)$	$\psi(x)$	$W(x)$	
HUBER				
If $\|x_i\| \le$ **c**	$\frac{x^2}{2}$	x	$\frac{c}{\|x\|}$	
If $\|x_i\| > c$	**c (|x | -** $\frac{c}{2}$ **)**	**X Sgn** (x)	**1**	**[6]**
TUKEY				
If $\|x\| \le$ **C**	$\frac{c^2}{6}(1-[1-(\frac{x}{c})^2]^2)$	$x[1-(\frac{x}{c})^2]^2$	$x[1-(\frac{x}{c})^2]^2$	
If $\|x\| >$ **C**	$\frac{c^2}{6}$			**[19]**
HAMPLE		$\psi(x)$		
		x ; **0 <** $\|x\|$ **< a**		
		a Sign(x**)** ; **a <** $\|x\|$ **< b**		
		a $\frac{(r-\|x\|}{r-b}$ **Sign (x)** ; **b≤** $\|x\|\le c$		
		0 ; **c <** $\|x\|$		**[6]**
Andrews	$\psi(x)$**= Sin(x)** ; $-\pi \le x \le \pi$			
	w_i**=** **sin (** $\frac{x_i}{c}$ **)/** $\left(\frac{x_i}{c}\right)$; $\|x_i\| \le c\pi$			
	0	$\|x_{i\|} > c\pi$		
	1			**[14]**

اذ ان : C= 4.685 ، Sign(x)= 1 ; $x > 0$ ، -1 ; $x < 0$

وتحسب معلمة القياس بنفس طريقة الامكان الاعظم :

$$s^{-1}f(\frac{x-\mu}{s})\ ;$$

$$\sum_i \psi\left(\frac{y_i-\mu}{s}\right)\left(\frac{y_i-\mu}{s}\right) = n$$

4-2 : (مقدرات $M-$) **الحصينة للنماذج المختلطة** ARMA(P, q)

لتكن النماذج المختلطة ARMA(P, q) طبيعية مستقرة، انعكاسية تمثل بالصيغة التالية :[12]

$$\emptyset(B)(X_t-\mu)=\theta(B)a_t$$
$$\emptyset(B)y_t=\theta(B)a_t+C$$

اذ ان : a_t **: متغيرات عشوائية تمثل الخطأ** ***(White noise)*** **تتوزع وفق توزيع طبيعي**

$\theta(B)\emptyset$ **(B),** : **عوامل متعددة الحدود جذورها** ***خارج حدود دائرة الوحدة***

$$\emptyset(B)=1-\sum_{i=1}^{p}\emptyset_i B^i$$
$$\theta(B)=1-\sum_{i=1}^{q}\theta_i B^i$$

وبفرض ان نموذج ARMA ملوث بجزأ من الخطأ $\in$

$$\lambda(B)=\emptyset^{-1}(B)\,\theta(B)=1+\sum_{i=1}^{\infty}\lambda_i B^i$$

وقد ذكر الباحث Huber "وجود شاردة واحدة في نموذج ARMA(P,q) يمكن ان يؤثر في جميع البواقي المتوافقة للفترة $t' \geq t$ **".**

تمثيل عمليات ***ARMA (P, q)*** **بوجود الشواذ بالصيغة التالية :**

X_t^{ϵ} = (1- ξ_t^{ϵ})$X_t + \xi_t^{\epsilon} w_t$

اذ ان :

X_t **نموذج :ARMA**

w_t **عملية اعتباطية :**

ξ_t^{ϵ} **: عملية تأخذ القيم (1,0)** ; $\lim_{n\to\infty}\frac{1}{n}(\sum \xi_t^{\epsilon}) = \epsilon$

في العمليات المستقرة $E(\xi_t^{\epsilon}) = \epsilon$

قياس

(تقدير- M) للخطأ a_t **يعد حل للمعادلة التالية**

$$E\left[\rho\left(\frac{a_t}{\sigma}\right)\right]=\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{U_i}{S}\right)\ ;\ U_i=u_1,u_2,\dots,u_n\in R$$

اذ ان :

$$\rho(x)=\rho(-x)$$

S: الانحراف المعياري ؛ $S \in 0, \infty()$

فيصبح النموذج الملوث تحت التأثير الأسي في المشاهدة اللاحقة ويظهر بفترات تذبذب قليلة على عكس ما هو في نموذج ARMA التقليدية .

$$y_t=\mu+a_t-\sigma_n\left(\frac{a_t}{\sigma}\right)+\sigma\,\emptyset^{-1}(B)\theta(B)n\left(\frac{a_t}{\sigma}\right)$$
$$=\mu+a_t+\sum_{i=1}^{p}\emptyset_1(y_{t-i}-\mu)-\sum_{i=1}^{r}(\emptyset_i a_{t-i}+(\theta_i-)\sigma_\eta\left(\frac{a_{t-i}}{\sigma}\right)$$

اذ ان : η **: دالة مفردة محددة بوجود k (تمثل حجم القيم الشاذة)**

R = Max (P, q) , $\theta > q = 0$, $\emptyset > p = 0$

σ : (مقياس- M) الحصين لـ a_t يمثل الانحراف المعياري بحالة التوزيع الطبيعي لـ a_t

وللحصول على مقياس $\left(\text{M} - \text{تقدير}\right)$ متسق مع الانحراف المعياري في بيانات طبيعية يجب تحقيق :

$$E\Phi(\rho(x)) = b$$

اذ ان :

Φ : تمثل التوزيع الطبيعي القياسي $N(0,1)$

S : الانحراف المعياري ; $b = \frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{U_i}{S}\right)$

وبتكرار $\left(M - \text{تقدير}\right)$ وفق التعديلات المبتكرة في نموذج ARMA نحصل على التوزيع الطبيعي التقاربي . ثم باستخدام اسلوب مونت كارلو للمقارنة بين تقدير ARMA الحصينة وقياس قوة حصانة التقدير باختيار التقدير الاصغر .

5-2 : مقدرات الامكان الاعظم (M-Estimate) للتوزيع الطبيعي :

The Normal Distribution Function:

يُعد التوزيع الطبيعي أحد أكثر التوزيعات المألوفة استخداماً في العمليات الإحصائية المختلفة. وعموماً فأن التوزيع يعرف بمعلمتي الموقع والقياس (μ ، S) على التوالي، وباستخدام الرموز السابقة يمكن كتابة دالة كثافة التوزيع وفقاً للصيغة الآتية: [19]

$$\text{Exp} = \frac{1}{2}\left(\frac{t-m}{s}\right)^2 \quad ; \quad S > 0$$

$$f\left(\frac{t}{M}, S\right) = \frac{1}{S\sqrt{2\pi}}$$

تقدير معلمات التوزيع بطريقة الامكان الاعظم MLE تتطلب تعظيم دالة Log-Likelihood

$$\text{Ln } L(M,\sigma^2) = \sum_{i=1}^{n} Ln\, f(x_i, \mu, \sigma^2)$$

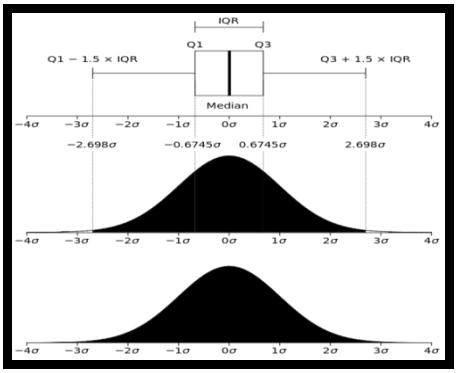
$$= -\frac{n}{2}\text{Ln}(2\pi) - \frac{n}{2}\text{Ln}\,\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x-\mu)^2$$

وبأخذ المشتقة بالنسبة الى μ ثم بالنسبة الى σ^2 وبحل المعادلات نحصل على تقديرات الامكان الاعظم :

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$S = \sqrt{Var(x_i)} \quad = \quad \text{معلمة القياس}$$



الشكل رقم (5) : معلمتي الموقع والقياس للتوزيع الطبيعي [19]

3- **الجانب التطبيقي** :

تم الحصول على قيم سلسلة زمنية متمثلة بالإشارة الكهربائية (Electrocardiogram -ECG) والتي تعكس تشخيص العوامل لمصادر التغير في السلسلة الزمنية ، وقد أعتمد على تسجيل الصورة البيانية لعمل القلب لشخص طبيعي (Normal subject) (غير المرضية). واعتمد نظام التسجيل على ما يعرف بالمتحسس السطحي الأحادي القطب والمكبر التماثلي الرقمي (AD 620)، بتمرير الإشارة من خلال المسيطر الدقيق (microcontroller) من النوع (ARDUINO) المتصل بجهاز الحاسوب ومن ثم معالجتها إحصائيا من خلال ما يتعلق بتحليل قدرة او قوة الطيف (Power Spectrum)،

وحيث أن طبيعة شكل الإشارة المبحوثة لا تتبع توزيعاً معيناً، الأمر الذي يجعل من الصعوبة اجراء عملية تقدير طيف القدرة في بعض الحالات بأتباع طرائق التقدير التقليدية التي تفرض توزيعاً إحصائيا محدداً، خاصة عندما يتعلق الأمر بالقيم المضافة (Additive Outliers-*AO*)، والذي يمكن لهذا النوع من التلوث أن يحجب الصفات المهمة للطيف، والثاني هو أن الإشارة المذكورة تتمثل بالتغيرات غير المنتظمة أو التي تعرف بطبيعتها العشوائية الحادة والتي تجعل النموذج الحقيقي المفترض لا يتناسب وطبيعة الظاهرة المبحوثة.. [1]

ولأجل تجاوز ذلك، جاءت فكرة الاعتماد على التقديرات الأولية الحصينة المعروفة بـ (M-Estimates) وهي طرائق التقدير الحصينة لـ (Huber، Tukey، Hampel، وطريقة التقدير الحصينة لـ Andrew) ، بهدف تقدير معلمتي الموقع والقياس للإشارة الحقيقية بدلاً من تعطيل أو تدمير عمليات النظام باستخدام اسلوب المحاكاة اعتمادا قيم التقديرات الأولية(Initial value) بشكل اعتباطي في ضوء نموذج إحصائي محدد، جاء استخدام تجارب التحليل هذه يعود الى طبيعة البيانات المبحوثة، من خلال ما يتعلق بتحليل قدرة او قوة الطيف (Power Spectrum)، أن هذا الأسلوب يُعد بمثابة عملية تنقية أو تنعيم للتقديرات الناتجة عن أتباع أساليب التقدير الإحصائية على البيانات الحقيقية والتي تكون نتائج التقدير فيها أكثر شمولاً أو تعميماً لمعلمات النموذج المفترض وقد استخدمها Harba [15]، من خلال توليده لأخطاء (White Noise) وتحميلها على خط الانحدار الموافق للبيانات المرتدة زمنياً وبافتراض التوزيع الطبيعي ، و أستخدمها Al-Naqeeb [79] في معالجة الإشارة الكهربائية للعضلة (EMG) في التحليل الطيفي الحصين .

3-1 : **فحص هوية القيم الشاذة**:

قبل البدء بتطبيق الطرائق المبحوثة لتحليل أطياف القدرة لا بد من أجراء الفحص الأولي للتعرف على هوية الإشارة الكهربائية (ECG)، حيث أن تشخيص نوع القيم الشاذة يرتبط بعملية الاختيار المناسبة لطريقة تقدير معلمات نموذج تحليل السلسلة الزمنية والتي يكون فيها التقديرات الحصينة هي الأنسب عندما يتعلق الأمر بنموذج القيم الشاذة من النوع المسمى بالقيم المضافة (Additive Outliers-*AO*)، وبناءً على ذلك لا بد من اختبار الفرضية الإحصائية الآتية:

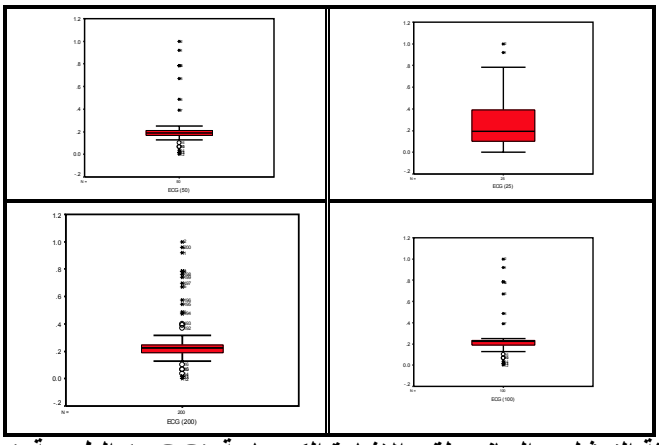
$$H_{\emptyset}: Model\ (IO)\ Hold$$
$$H_1: Model\ (AO)\ Hold$$

بعد تسجيل البيانات لقيم الاشارة (ECG) على تطبيق EXCEL يتم نقل المقاطع الاربعة الى تطبيق SPSS لفحصها اذا كانت تتضمن شواذ من النوع القيم التجميعية (Additive Outliers-AO) والتي تبعد عن قيمة المتوسط ب 3

SPSS → ANALYSIS → DISCRIBTIVE → EXPLORER → STATISTICS → (M – ESTIMATE)

الجدول(1): نتائج إحصاءه اختبار تشخيص نوعي القيم الشاذة للإشارة الحقيقية وللمقاطع الزمنية المختارة ومستوى الدلالة 0.000

سعة المقطع الزمني(ms)	إحصاءه الاختبار	نوع القيم الشاذة
25	-3.431	*AO*
50	-8.418	*AO*
100	-17.259	*AO*
200	-25.167	*AO*



الشكل رقم (6): طريقة التمثيل – البياني لقيم الإشارة الكهربائية (ECG) الطبيعية (عدم الإجهاد) وللمقاطع الزمنية المختارة بـ (ms).

2-3: **تقدير قيم المعلمات (الموقع والقياس)**:

ولكافة المقاطع الزمنية المختارة، بهدف أمكانية تنفيذ تجارب التنعيم بالقيم الابتدائية لمعلمات النماذج الإحصائية المدروسة لقيم الإشارة (ECG) لابد من تقدير معلمتي الموقع والقياس ووفقاً للمقاطع الزمنية المختارة، وتقدير القيم الابتدائية بالطرائق الحصينة المعروفة بـ (M- Estimates) الحصينة لـ (Huber، Tukey، Hampel، Andrew، MLE) المعلمية، وبعد عمليات تنعيم بيانات الاشارة الكهربائية، نكرر العمليات السابقة باستخدام الحزم الاحصائية spss وبرنامج excel في ايجاد معلمتي الموقع والقياس(m-Estimate ووفق التوزيع المدروس .

الجدول (3-2-1): التقديرات الحصينة و المعلمية لمعلمتي الموقع والقياس للبيانات الحقيقية

تقديرات القيم الابتدائية - (للإشارة الحقيقية)		المقاطع الزمنية (ms)	طريقة التقدير	Statistics
Scale	Location			
0.109361071	0.198240761	25	Huber	((Robust))
0.110396146	0.152001176	25	Tukey	
0.110042114	0.192264205	25	Hampel	
0.115699038	0.150987823	25	Andrew	
0.299104021	0.294758479	25	MLE	Parametric
0.110547941	0.189450318	50	Huber	((Robust))
0.110568877	0.189450318	50	Tukey	
0.110934697	0.186333709	50	Hampel	
0.110561124	0.189341268	50	Andrew	
0.216218575	0.242587022	50	MLE	Parametric
0.108189273	0.301032780	100	Huber	((Robust))
0.108067007	0.296113961	100	Tukey	
0.108293787	0.299006545	100	Hampel	
0.108068333	0.296079735	100	Andrew	
0.152303878	0.237134980	100	MLE	Parametric
0.145088439	0.218847828	200	Huber	((Robust))
0.145671924	0.215956964	200	Tukey	
0.145866335	0.215050158	200	Hampel	
0.145662682	0.216000714	200	Andrew	
0.142396157	0.246598862	200	MLE	Parametric

3-3: التقدير الطيفي لقيم الإشارة :

يتم توليد البيانات العشوائية وفق اسلوب المحاكاة المقيدة وبموجب التطبيق العشوائي (***Stat graphics***) ثم تنقل من خلال الايعاز ***Edit*** وتلصق بملف خاص على التطبيق الجاهز ***Excel***.وبالخطوات التالية :

Describe ⇢ *Distribution* ⇢ (*Probability Distribution*)

بهدف تقدير قيم مربع السعة (Amplitude) في التحليل الطيفي لقيم الإشارة الكهربائية (ECG) المحاكاة بعلاقتها بالتردد (Frequency) والتي تعرف بطيف القدرة (Power Spectrum) وهي موضوع التقديرات الحصينة لاحتواء قيم الإشارة على صنف القيم الشاذة المضافة (*AO*) ولكافة المقاطع الزمنية المبحوثة، من خلال استخدام الطرائق الحصينة المعروفة بـ (M- Estimates) وهي (طريقة التقدير الحصينة لـ (Huber)، (Tukey)، (Hampel)، وطريقة التقدير الحصينة لـ (Andrew) بالاضافة الى طريقة MLE المعلمية وللمقاطع الزمنية المختارة ، بغية أمكانية أجراء المقارنات ما بين نتائج أطياف القدرة باختلاف طرائق التقدير من جهة، والمقطع الزمني للإشارة، فقد تم تقدير نسبة سعة أشارة الطيف الى الضوضاء (μ/S^2) عند كل مقطع، والممثلة للقياس غير المباشر لسرعة انتقال الإشارة (ECG).

4-3: التقديرات الحصينة لأطياف القدرة بموجب النموذج (الطبيعي):

تنقل نتائج المقطع بأكمله الى التطبيق NCSS *لأجراء التحليل الطيفي وتقدير معلمات النموذج من خلال التطبيق* NCSS *وفق الخطوات التالية :*

Analyze → *Forecasting Time Series* → *Spectral Analysis* → *OK*

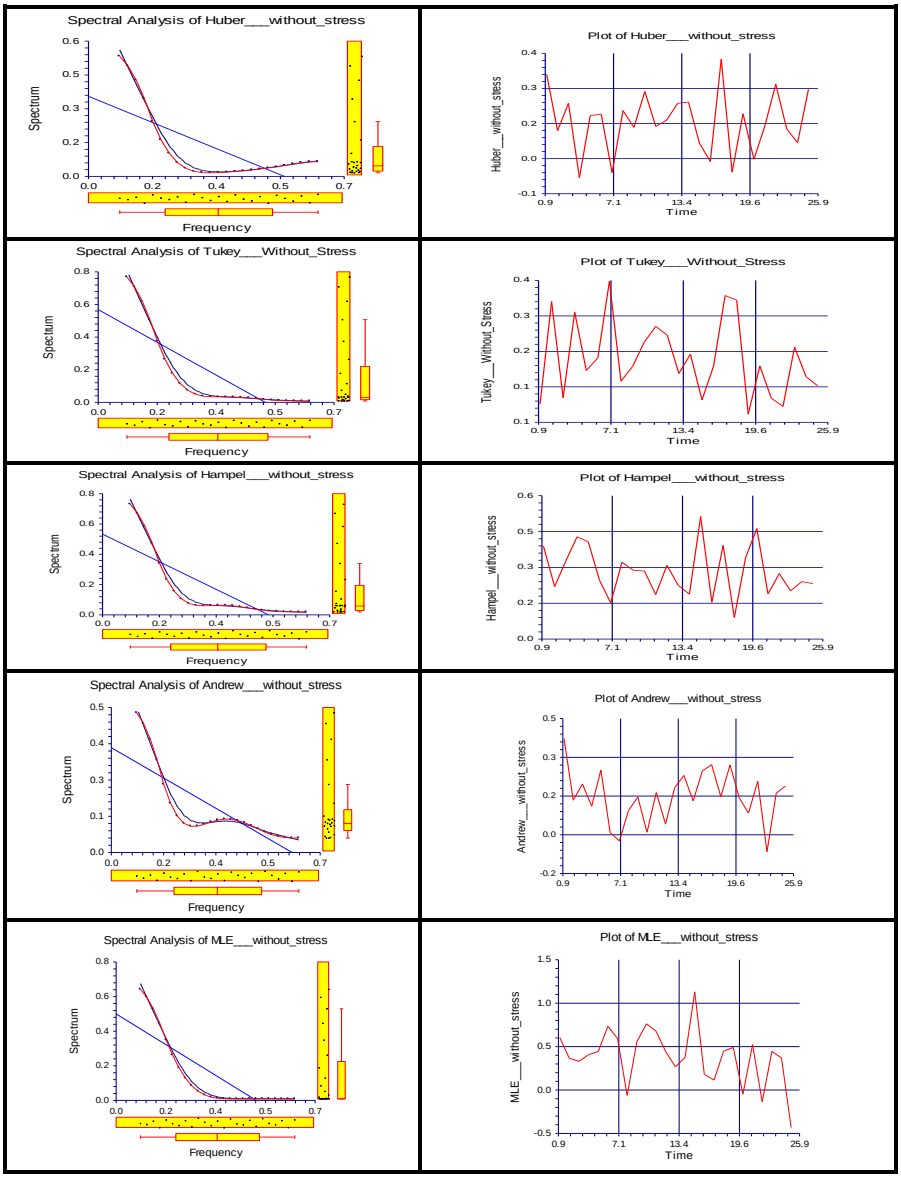
الجدول (3-4-1): التقديرات الحصينة لمكوني نسبة الطيف الى الضوضاء الحصينة للإشارة الكهربائية (ECG) لبيانات لتوزيع الطبيعي المنعمة ولكافة المقاطع الزمنية المختارة

التقديرات الحصينة لأطياف القدرة - (للإشارة المحاكاة)			المقاطع الزمنية (ms)	طريقة التقدير	Statistics
μ/S^2	(S^2)	μ			
59.78731005	0.019808959	1.18432438	25	Huber	((Robust))
19.59624739	0.035707685	0.69973663	25	Tukey	
112.1901123	0.028631019	3.21211720	25	Hampel	
105.680153	0.019787373	2.09113262	25	Andrew	
99.00686636	0.123318072	12.2093359	25	MLE	Parametric
214.3719349	0.012268819	2.63009056	50	Huber	((Robust))
162.5951296	0.015218660	2.47447999	50	Tukey	
227.1939055	0.016489494	3.74631250	50	Hampel	
166.8249176	0.016484008	2.74994328	50	Andrew	
198.0330532	0.060881813	12.0566114	50	MLE	Parametric
1264.289253	0.019070755	24.1109508	100	Huber	((Robust))
1270.885397	0.016908479	21.4887389	100	Tukey	
653.1554869	0.023662636	15.4553806	100	Hampel	
826.0778034	0.023143735	19.1185260	100	Andrew	
618.5853948	0.047280800	29.2472125	100	MLE	Parametric
2112.554021	0.028346273	59.8830334	200	Huber	((Robust))
2693.414890	0.034016493	91.6205291	200	Tukey	
1985.793425	0.037293918	74.0580173	200	Hampel	
1807.805375	0.032488051	58.7320731	200	Andrew	
2545.633711	0.038356987	97.6428404	200	MLE	Parametric

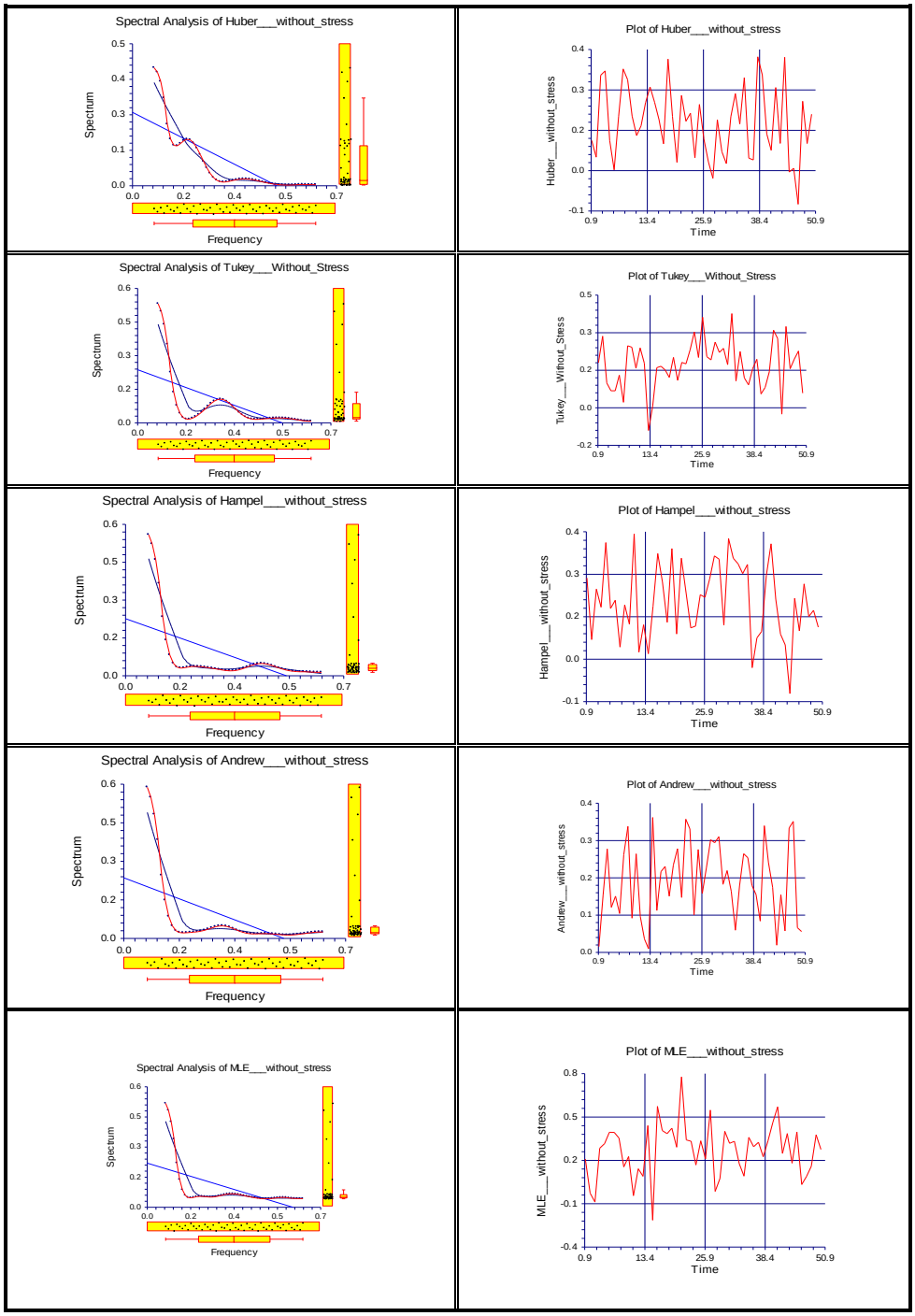
حيث تشير نتائج تقديرات نسبة الطيف الى الضوضاء وفقاً للمقدرات الحصينة الى ارتفاع درجة الاختلاف ما بين الطرائق الحصينة من جهة، والى الارتفاع المميز في قيم نسبة الطيف الى الضوضاء بموجب الطريقة التقليدية بالتوليد المعلمي بموجب طريقة MLE وعند كافة المقاطع الزمنية المختارة للإشارة (ECG) عموماً.

والأشكال البيانية رقم (3-4-1-1)، (3-4-1-2)، (3-4—1-3)، و (3-4-1-4) توضح مخططات قيم الإشارة (ECG) المحاكاة بالطريقة المقيدة أضافة الى طيف القدرة للمقاطع الزمنية المحددة بـ (TSV= 25 ms، TSV= 50 ms، TSV= 100 ms، و TSV= 200 ms) على التوالي بافتراض التوزيع الطبيعي.

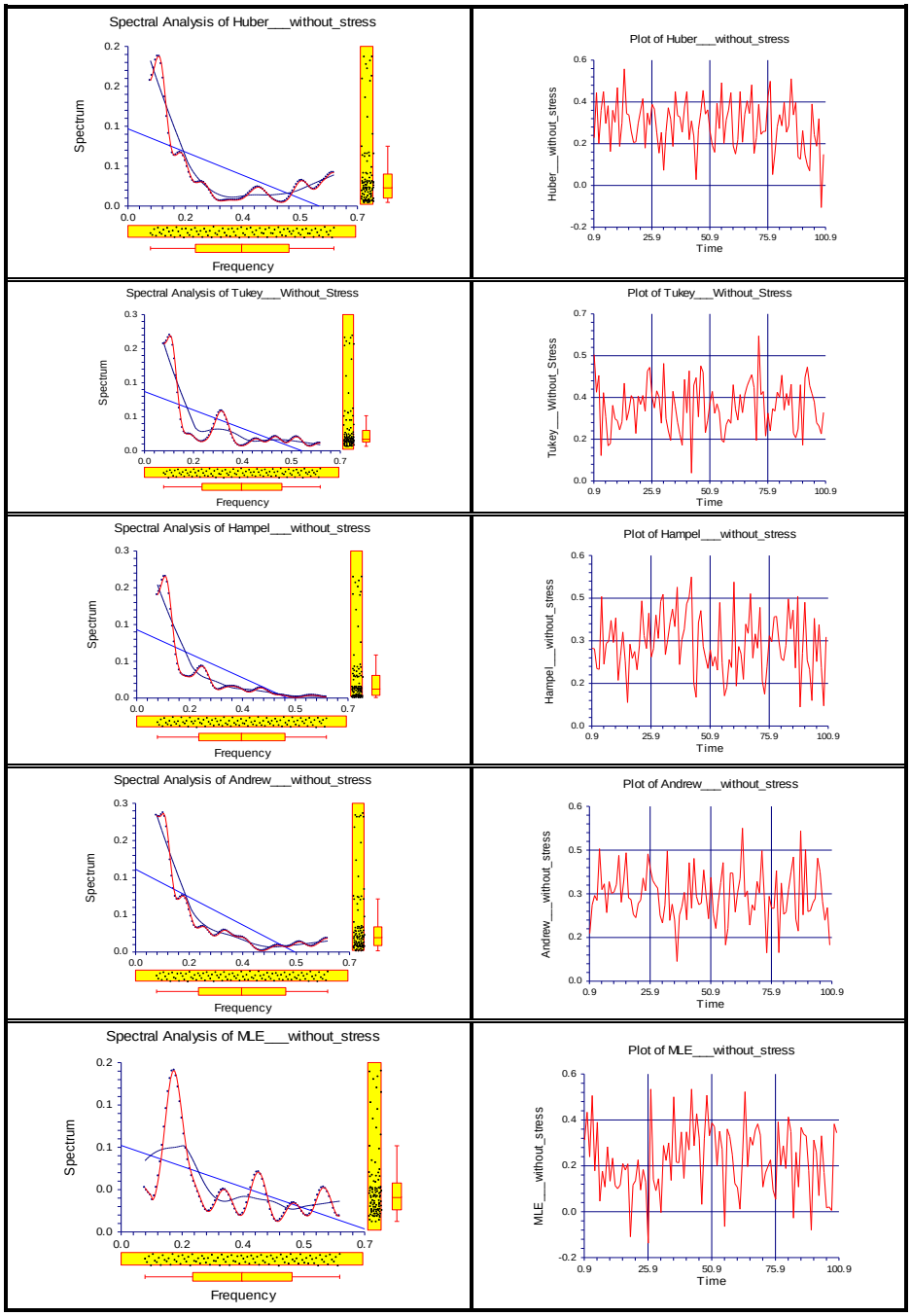
الشكل (3-4-1-1): مخطط قيم الإشارة (ECG) بالمحاكاة المقيدة و طيف القدرة للمقطع الزمني بافتراض التوزيع الطبيعي (TSV= 25 ms)



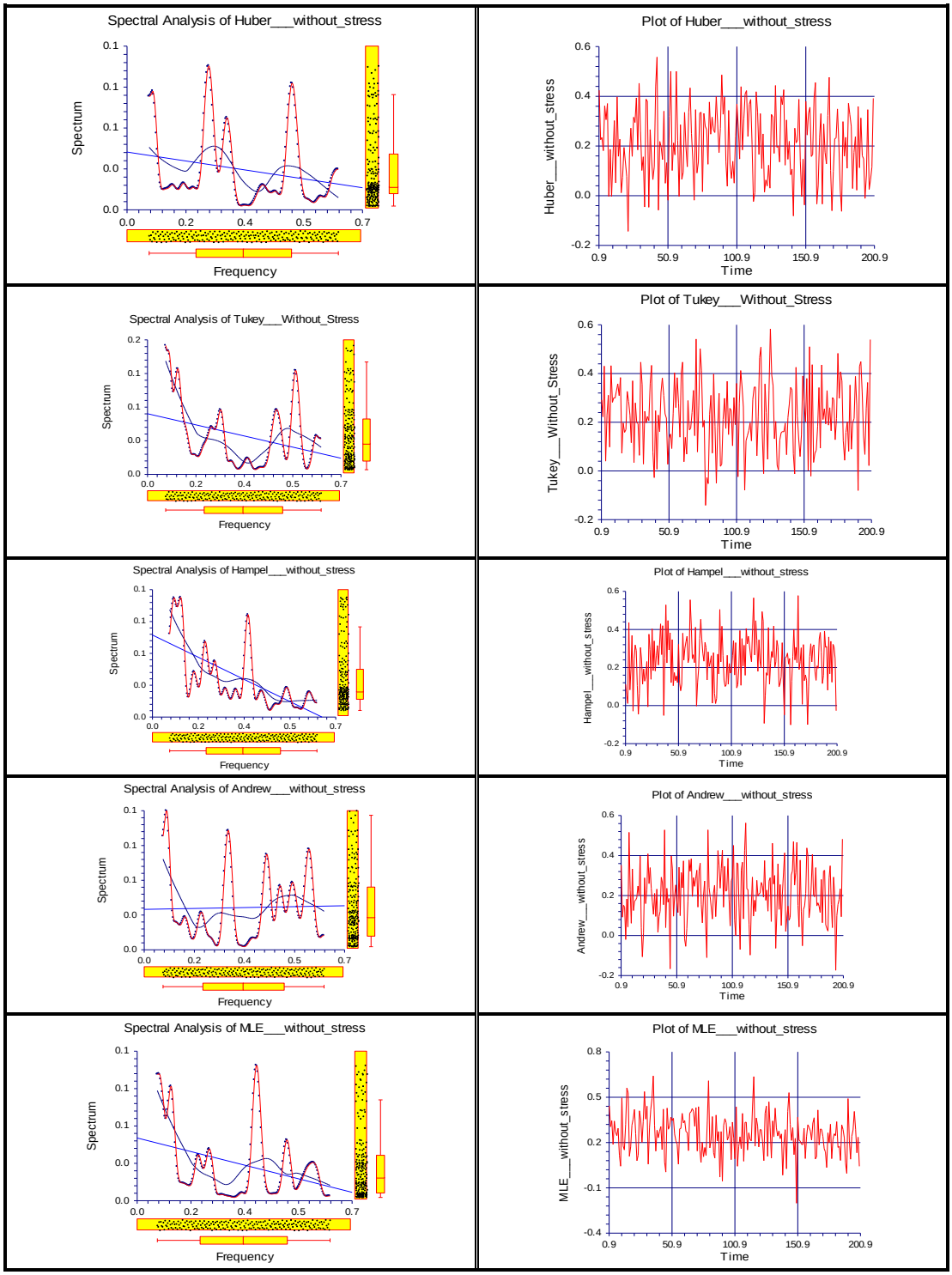
الشكل (4-4-1-2): مخطط قيم الإشارة (ECG) و طيف القدرة للمقطع الزمني (TSV= 50ms) بافتراض التوزيع الطبيعي



الشكل (4-4-1-3): مخطط قيم الإشارة (ECG) و طيف القدرة للمقطع الزمني (TSV= 100 ms) بافتراض التوزيع الطبيعي



الشكل (4-4-1-4): مخطط قيم الإشارة (ECG) و طيف القدرة للمقطع الزمني (TSV= 200 ms) بافتراض التوزيع الطبيعي



الاستنتاجات :

1.: لقد حقق الأسلوب المتبع في تنقية أشارة الضوضاء البيضاء بتقييد معلمتي الموضع والقياس من خلال اعتماد التقديرات الأولية الناتجة عن كافة التجارب بتطبيق طرائق التقدير الحصينة لمعلمتي الموقع والقياس على الإشارة الحقيقية في الحصول على نتائج واقعية ترتبط بالنظام الحقيقي لدراسة سلوك الإشارة الكهربائية (ECG) وذلك نتيجة لاحتواء قيم الإشارة على الملوثات بالقيم المضافة.
2.أن اعتماد مؤشر الحالة الغير مباشرة لسرعة انتقال الإشارة في أجراء عمليات المفاضلة وتقدير معلمات النموذج المختلط من خلال بيان مدى صلاحية طرائق التقدير المستخدمة قد تحدد بالأساس على نتائج تقديرات تباينات أخطاء الابيضاض الممهدة بشكل خاص، وبالرغم من أتساع أشارة طيف القدرة يعكس الصفة التباعديه (Divergent) للتوزيع الملوث، الأمر الذي يؤكد مستوى التنقية والتنعيم لإشارة الضوضاء البيضاء الممهدة بموجب طريقة التقدير المحددة .

التوصيات :

لأهمية استخدام النموذج المختلط ARMA لتقديرات قيم أطياف القدرة بموجب الطرائق الحصينة ووفق النماذج الإحصائية المختلفة، فأنه يمكن التوسع باستخدام نماذج أخرى للتقدير قد تقتضيها طبيعة الظاهر موضوع البحث، أمثال نموذج الانتقال العشوائي (Random Walk)، النموذج التربيعي لمتعدد الحدود (Quadratic Trend)، النموذج الأسي (Exponential Trend)

المصـــادر :

1- Abdul Khaleq A. Al-Naqeeb, Ayad A. Ibraham, Faez F. Shareef,(2015) " Power Spectra Robust Estimation of (EMG) Signals Using Autoregressive Simulation", Proceeding of the World Congress on Electrical Engineering and Computer Systems and Science (EECSS) Barcelons, Spain, No.294 .
2- AL-Naqeeb, A. A., & AL- Doulaimi, M.M.A. (1997). "Robust Estimation of Auto-Regressive Processor AR(P) and the Optimal Using in the Artificial Electronic Limbs", Administrated and Economic Journal University of Baghdad. Baghdad Iraq.
3- Ben-Gal Irad, Maimon O. And Rockach L.(2005) و " Data Mining And Knowledge Discovery Handbook: A Complete Guide For Practitioners And Researchers - CH_1- Outlier Detection", Kluwer Academic Publishers, ISBN 0-387-24435-2.
4- Daniel PEÑA and Victor YOHAI,(1999) " A Fast Procedure for Outlier Diagnostics in Large Regression Problems", American Statistical Association, Journal of the American Statistical Association , Vol. 94, No. 446, Theory and Methods.
5- Djuri´ C Petar M.,& Kay S.M.(1999)," Spectrum Estimation And Modeling", Digital Signal Processing Handbook-Ch14,State University Of New York At Stony Brook , University Of Rhode 15) Island.
6- Frank Hampel,(2001)," Robust statistics: A brief introduction and Overview, Seminar for Statistic , Eidgenossische Technische Hochschule (ETH) CH-8092 Zurich Switzerland.
7 - Huber, P.J.,(1980)," Robust Regression Asymptotic Conjectures and Monte-Carlo";Ann.Statist.1,pp799-821

8- I.A. Mcleod ,(1978)," On the distribution of Residual Autocorrelation in Box-Jenkins Models", J. R. Statist .Soc. B 40,No. 3,pp 296-302.
9- Nora Muler[1], Daniel Pe˜na[2] And V´Ictor J. Yohai[3],(2009)," Robust Estimation For ARMA Models" , Universidad Torcuato Di Tella, Universidad Carlos Iii De Madrid And Universidad De Buenos Aires And Conicet , Vol. 37, No. 2, 816–840, Institute Of Mathematical Statistics.
10- Rand R. Wilcox, H. J. Keselman,(2003)," Modern Robust Data Analysis Methods: Measures of Central Tendency", University of Southern California[1], University of Manitoba[2] , Psychological Methods Copyright 2003 by the American Psychological Association, Inc., Vol. 8, No. 3, 254–274.
11- Rudolf Dutter ,(1987),"Robust Time Series Analysis: A Survey" , Norbert Stockinger. O. University Journal Kybernetika,vol.23. 52) Sakia, R. M.," The Box-Cox transformation technique: a review", The Statistician (1992) 41, pp. 169-178.
12- S. B. L. Prakasa Rao,(1979)," A Asymptotic Behavior of M- estimators for the Linear Model With Dependent Errors", Purdue University, Indian Statistical Institute, New Delhi.
13- Taylor Jonathan,(2003)," Introduction to Regression and Analysis of Variance- Robust methods"
14- T.W. Anderson,(1975),"Maximum Likelihood Estimation Of Parameters of Autoregressive Processes With Moving Average Residuals and Other Covariance Matrices With Linear Structure" The Annals of Statistics , Vol.3, No. 6. 1283-1304.
15- M. I Harba, A., Lynn, P. A., 1981. "Optimizing the acquisition and processing of surface EMG signals" J. Biomed. Eng. Vol. 3. 100-106.
16- Narel Hausitoe, Daniel Maposa2 And 'Maseka Lesaoana3(2012)," A Method For Detection And Correction Of Outliers In Time Series Data",1department Of Applied Mathematics, National University Of Science And Technology, African Journal Of Business Management Vol.6 (22), Pp. 6631-6639.
17- P.J.Huber, ,(1980)," Robust Regression Asymptotic Conjectures and Monte-Carlo";Ann.Statist.1,pp799-821.
18- Qaren Kafadar,(2001)," John Tukey and Robustness", Department of Mathematics, University of Colorado-Denver, Denver, Colorado, American Statistical Association, 80217-3364.
19- Stefan Mittnik,"On Efficient Exact Maximum Likelihood Estimation of High-Order ARMA Models", Department of Economics, State University of New York at Stony Brook, Stony Brook, New York 11794-4384.

Some Robust methods for Estimates the power Spectrum in ARMA Models Simulation Study

Abstract :

Robust statistics Known as, resistance to errors caused by deviation from the stability hypotheses of the statistical operations (Reasonable, Approximately Met, Asymptotically Unbiased, Reasonably Small Bias, Efficient) in the data selected in a wide range of probability distributions whether they follow a normal distribution or a mixture of other distributions deviations different standard .

power spectrum function lead to, President role in the analysis of Stationary random processes, form stable random variables organized according to time, may be discrete random variables or continuous. It can be described by measuring its total capacity as function in frequency.

As supposed distribution of the data mistakes in accordance with the normal or approach to the normal distribution، Share estimation methods fortified with the concept of nonparametric in the absence of a clear model parameters (Free distribution) as the contamination is unknown model has distributed part of it in accordance with the normal distribution, while the other part is unknown distribution, and thus became the distribution of tainted its parameters unknown, so it can be considered the highest level in grades nonparametric methods are based on the conversion calculable test to a standard degree held by the convergence class standard process.

The aim of the Search: compared to some of the Robust estimation methods and non-parametric, So find the best estimator of Power spectrum With the mixed ARMA model for time series follow a Normal distribution, Then, applied the most accurate on realistic data About the diagnosis of the factors that reflect the sources of change in time series of the electrical reference (ECG - Electrocardiogram) in normal state (unsatisfactory), through the analysis of the power spectrum. By finding the best estimator of the ability of spectrum according to the mixed spectrum for Series follow a particular dist. & applied to Real data(ECG), The method adopted other Parameters to get realistic results are related to the behavior of the real System reference (ECG)the Capabilities of ability Normal dist. Increase with the size of the Section ,and the highest at Section 200 in a manner Tokey.

Key Word : **Robust, Power Spectrum, ARMA, Simulation**